A new symptom model for autism cross-validated in an independent sample

Boomsma, A.; Van Lang, N. D. J.; De Jonge, M. V.; De Bildt, A. A.; Van Engeland, H.; Minderaa, R. B.

Published in:
Journal of Child Psychology and Psychiatry

DOI:
10.1111/j.1469-7610.2008.01897.x

Publication date:
2008

THE JOURNAL OF
CHILD PSYCHOLOGY AND PSYCHIATRY
ACAMH
Journal of Child Psychology and Psychiatry 49:8 (2008), pp 809–816 doi:10.1111/j.1469-7610.2008.01897.x

A new symptom model for autism cross-validated in an independent sample

A. Boomsma,[1] N.D.J. Van Lang,[2] M.V. De Jonge,[3] A.A. De Bildt,[4] H. Van Engeland,[3] and R.B. Minderaa[4]

[1]University of Groningen, Department of Statistics and Measurement Theory, The Netherlands; [2]Leiden University Medical Center, Curium, Department of Child and Adolescent Psychiatry, The Netherlands; [3]University Medical Center Utrecht, Department of Child and Adolescent Psychiatry, The Netherlands; [4]University Medical Center Groningen, Department of Child and Adolescent Psychiatry, The Netherlands

Background: Results from several studies indicated that a symptom model other than the DSM triad might better describe symptom domains of autism. The present study focused on a) investigating the stability of a new symptom model for autism by cross-validating it in an independent sample and b) examining the invariance of the model regarding three covariates: symptom severity, intelligence, and age. **Method:** The validity of the symptom model was examined in an independent sample of *N = 263* children and adolescents with autism spectrum disorders, and model invariance was studied in a larger sample of *N = 356* children and adolescents with autism spectrum disorders. The fit of the symptom model to the sample data was compared to that of alternative models (including the DSM triad), and the invariance of the new model was investigated for each covariate by multiple-group comparisons. **Results:** The fit of the new symptom model was better than that of two alternative models. It could not be compared to that of the DSM triad, because the latter encountered empirical identification problems. There were no significant or substantive differences between the estimated model in each of the dichotomised groups for any of the three covariates, which indicated factorial invariance of both structural form and factor loadings. **Conclusions:** The symptom model appeared to be relatively stable: It could be cross-validated in the independent sample and factorial invariance was shown between the dichotomised groups for each covariate. Further model validation with instruments other than the Autism Diagnostic Interview-Revised (ADI-R) is recommended. **Keywords:** Autism spectrum disorder, symptom model, cross-validation.

Several studies investigated symptom change and symptom structure in autism, as related to the three symptom domains of the DSM-IV, i.e., impaired social interaction, impaired communication, and the occurrence of stereotyped behaviours and restricted interests (American Psychiatric Association, 2000). The results of these studies indicated that the symptoms of autism in the three domains develop differently over time, and that a symptom structure other than that of the DSM might be more appropriate to comprehend symptom change.

Some longitudinal studies, for instance, demonstrated that the diagnoses of autism spectrum disorders (ASD) remain fairly stable over time, whereas the symptoms of the three DSM domains develop differently. Charman et al. (2005), for example, found that symptoms of impaired nonverbal communication improved each year in young childhood (i.e., from age 2 to age 7), while symptoms of impaired social interaction improved from age 4–5 years on, and symptoms of stereotyped behaviours and restricted interests worsened before they improved. In other studies it was demonstrated that language skills related to impaired communication developed most strongly in the early school years but not between middle childhood and late adolescence (Sigman & McGovern, 2005), and symptoms of impaired social interaction and stereotyped behaviours and restricted interests improved between middle-childhood and late adolescence (McGovern & Sigman, 2005).

Apart from the issue of symptom change differences, two studies by Ronald et al. (2006a, 2006b) revealed that the three DSM domains are relatively separate constructs. They found that the correlations between the three symptom domains were modest (ranging between .10 and .40) in more than 3,000 twin pairs from a general population (Ronald et al., 2006a; Ronald, Happé, Price, Baron-Cohen, & Plomin, 2006b). They also showed that these correlations remained modest in a subgroup of twin pairs having relatively extreme autistic scores.

These findings indicate that the three DSM domains are relatively independent and that change in autistic symptoms should be evaluated in the three domains separately. Beside the relative independence of the DSM symptom domains, there is support for a different composition of symptoms within the domains. Kolevzon, Smith, Schmeidler, Buxbaum, and Silverman (2004) showed that within the DSM domains, symptoms may have a strong or weak familiality. They conducted a family study of 16 monozygotic siblings, and found a significant familiality in symptoms associated with impaired

Conflict of interest statement: No conflicts declared.

Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA

socialisation and impaired communication, but not with stereotyped behaviours. They found more differences in familiality when specific symptoms within the DSM symptom domains were considered, i.e., subdomain scores instead of domain scores in the algorithm of the Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter, & Le Couteur, 1994). It was shown that social play, which is incorporated in the DSM domain impaired communication, had no familiality, whereas circumscribed interest and preoccupation with sensory items or parts of objects, both from the DSM stereotyped behaviours domain, had significant familiality.

The results from a number of other studies suggested also a composition of the symptom domains other than that of the DSM. It was shown, for instance, that symptoms of the DSM domains impaired social interaction and impaired communication may represent one symptom domain to be described as impaired joint attention and affective reciprocity (Robertson, Tanguay, L'Ecuyer, Sims, & Waltrip, 1999; Tanguay, Robertson, & Derrick, 1998), impaired social intent (Tadevosyan-Leyfer et al., 2003), or impaired social communication (Georgiades et al., 2007; Van Lang et al., 2006). It has further been shown that symptoms such as particular speech characteristics and behavioural features, like compulsions and restricted interests, are related (Tadevosyan-Leyfer et al., 2003; Van Lang et al., 2006); and that several symptoms from the DSM domains about difficulties in play skills are associated (Robertson et al., 1999; Van Lang et al., 2006).

Van Lang et al. (2006) proposed a new symptom model for autism that may account for some of the findings of Kolevzon et al. (2004). This new model consists of three main symptom factors: 1) The factor *impaired social communication* is composed primarily of symptoms from DSM impaired social interaction, but also of symptoms concerning inadequate use of gestures and failure to initiate or sustain conversational interchange from DSM impaired communication. 2) The factor *stereotyped features in speech and behaviour* is reflected by symptoms from DSM stereotyped behaviours and restricted interests, but it also comprises symptoms from DSM impaired communication, like idiosyncratic or repetitive speech characteristics. 3) The factor *impaired play skills* is composed of symptoms from DSM domains impaired social interaction and impaired communication, i.e., failure to develop peer relationships and lack of play skills.

Confirmatory analysis corroborated this three-factor symptom model in a group of 255 verbal and nonverbal children and adolescents with minor to severe autistic symptoms and a full-scale intelligence score (FIQ score) larger than 20 (see Van Lang et al., 2006). Given the available sample data, however, hypothesized effects of symptom severity, intelligence, and age on the symptom model could not be examined thoroughly. Symptom severity, intelligence, and age are important variables in describing autism heterogeneity (e.g., Coplan & Jawad, 2005). Therefore, to further substantiate the new symptom model, the present study was conducted to examine: a) the model's validity in a new and independent sample of ASD children and adolescents, and b) the model's stability and invariance in dichotomised groups of ASD children and adolescents with regard to symptom severity, intelligence, and age.

Methods

Sample

In the current study, two Dutch samples were used. For the cross-validation part, data from a sample collected by the University Medical Centre Utrecht were employed. For the investigation of the invariance of the symptom model under dichotomised covariate effects, sample data from Utrecht were combined with those from the University Medical Centre Groningen; the symptom model was initially constructed and evaluated using the Groningen sample data. The following inclusion criteria were applied to both samples: all participants had a clinical disorder on the autism spectrum, were verbal (i.e., functionally using at least three word phrases according to the ADI-R), had an FIQ score ≥35, and were aged between 4 and 25 years at the time the ADI-R was administered. Parents of all participants gave written informed consent.

Data from Utrecht. *N = 263* ASD children and adolescents were selected from a large data set of *N = 374* individuals who participated in the ongoing collaborative genetic study of autism (International Molecular Genetic Study of Autism Consortium (IMGSAC), 2001) or in other imaging or medication studies in Utrecht. These data sets could be combined because the parents of all 374 participants were interviewed with the ADI-R. The participants were recruited from outpatient and inpatient clinics of the Department of Child and Adolescent Psychiatry from the University Medical Centre in Utrecht, from other specialised centres for autism, or in cooperation with the Dutch parent association for autistic children. The clinical diagnosis of each participant was established by an experienced clinician who studied the medical records, the developmental history and available diagnostic information, like ratings on the ADI-R and Autism Diagnostic Observation Schedule (ADOS-G; Lord et al., 2000). From the 374 participants, 111 were excluded because they did not meet the inclusion criteria or because the time between the assessment of the ADI-R and IQ was too large (i.e., more than three years) to ensure that the IQ measurement was reliable at the time of the ADI-R interview. More specifically: 102 were excluded because the IQ assessment date was missing, or because FIQ scores were considered unreliable, and 9 were excluded because they were nonverbal (8) or aged ≥26 (1). As this study aimed at examining symptom model invariance for different IQ levels, it was decided to include only participants with IQ data obtained from psychometrically sound instruments, like the Wechsler scales (Van der

Steene et al., 1986), Raven Progressive Matrices (Van Bon, 1984) and the Dutch intelligence test RAKIT (Bleichrodt, Resing, Drenth, & Zaal, 1987).

Data from Groningen. *N = 93* ASD children and adolescents were selected for this study. The original Groningen sample consisted of 308 participants; the symptom model was previously tested in a selection of 255 verbal and nonverbal participants with minor to severe autistic behaviours, an FIQ score >20, and aged between 4 and 20 years. For details on the primary data collection in Groningen, see De Bildt, Sytema, Kraijer, and Minderaa (2005), and Van Lang et al. (2006). In the present study, 162 participants from these 255 were excluded, because they did not meet the inclusion criteria: 85 did not fulfil the criteria for a disorder on the autism spectrum, 66 had FIQ score <35, 25 of whom were also nonverbal, and 11 were nonverbal with a moderate (7) or mild (3) FIQ, or a FIQ in the borderline range (1). Most of them (125) were typically developing, 16 had a clinical diagnosis of Pervasive Developmental Disorder – Not Otherwise Specified (PDD-NOS) and 21 of Autistic Disorder (AD).

The main characteristics of the sample data from Utrecht and Groningen are presented in Table 1.

Latent variable indicators and model descriptions

The symptom factor model under investigation is denoted as Model B2 (cf. Van Lang et al., 2006). Its measurements are the 12 subdomain scores from the ADI-R algorithm that refer to the three DSM symptom domains: impaired socialisation (*S*), impaired communication (*C*), and stereotyped behaviours and restricted interests (*R*). These 12 ADI-R subdomain scores, here defined as ADI-R indicators, represent sumscores of 2 to 4 items (for details about these items and subdomain scores, and output of the major analyses, see the website http://www.gmw.rug.nl/~boomsma/lang2.htm). As indicated in the ADI-R algorithm, item scores of 3 were recoded as 2.

The 12 ADI-R indicators are described as follows: failure to use nonverbal behaviours to regulate social interaction (*S1*), failure to develop peer relationships (*S2*), lack of shared enjoyment (*S3*), lack of socio-emotional reciprocity (*S4*), lack of, or delay in, spoken language and failure to compensate through gesture (*C1*), lack of varied spontaneous make-believe or social imitative play (*C2*), relative failure to initiate or sustain conversational interchange (*C3*), stereotyped, repetitive or idiosyncratic speech (*C4*), encompassing preoccupation or circumscribed pattern of interests (*R1*), compulsive adherence to non-functional routines or rituals (*R2*), stereotyped and repetitive motor mannerisms (*R3*), preoccupations with part-objects or non-functional elements of materials (*R4*).

For the cross-validation, the three-factor Model B2 (factor *F1* with *S1*, *S3*, *S4*, *C1* and *C3*; factor *F2* with *C4*, *R1*, *R2*, *R3* and *R4*; and factor *F3* with *S2* and *C2*, as indicators) was tested in the sample of Utrecht. As had been done previously for the Groningen data (Van Lang et al., 2006), for the Utrecht data it was also investigated whether Model B2 fitted better than the DSM model, or a one- or two-factor model. The DSM model consisted of three factors (a factor with *S1* thru *S4*; a factor with *C1* thru *C4*; and a factor with *R1* thru *R4*, as indicators). In the one-factor model all indicators *S1* thru *R4* referred to one single autism factor. The two-factor model had a factor with *S1* thru *C4*, and a second one with *R1* thru *R4*, as indicators. All four models were tested and evaluated using a) ADI-R ratings related to the *current status*, and b) ADI-R ratings related to the most abnormal age period of 4–5 years, or ever, the *past status*.

For the investigation of possible effects of the three covariates on the symptom model, Model B2 was tested in selected groups from the joint samples of Utrecht and Groningen. Within that combined sample (*N = 356*), three selections were made to investigate the invariance of the model structure and factor loadings: a) the effect of symptom severity was examined by comparing Model B2 estimates in ASD participants having autism (AD; *N = 143*) or PDD-NOS or Asperger Syndrome (PDD-NOS/AS; *N = 213*). b) The effect of intelligence was studied by comparing Model B2 estimates for participants having an FIQ score <85 (*N = 136*) versus an FIQ score ≥85 (*N = 220*). c) The effect of age was examined for children of age 4–12 (*N = 247*) versus adolescents of age 13–24 (*N = 108*). The overlap between the dichotomised scores on the three covariates in the selected groups was modest: the largest overlap was found for 97 children who had PDD-NOS/AS and a high FIQ score. Here, only results regarding the ADI-R *current status* are presented, because it was decided that these ratings reflect participant's behaviour more

Table 1 Sample characteristics

	Data from Utrecht	Data from Groningen
N	263	93
Male : Female	227 : 36	79 : 14
Age	4–24 years (mean 11 yrs; SD 5 yrs)	4–19 years (mean 11 yrs; SD 4 yrs)
FIQ	35–152 (6% with FIQ < 70)	35–129 (63% with FIQ < 70)
Clinical diagnosis of AD[1]	110 (42%)	33 (35%)
Clinical diagnosis of PDDNOS/AS[2]	153 (58%)	60 (65%)
Mean score (SD) on the ADI-R domains:		
Impaired social interaction	19.1 (6.5)	19.2 (6.2)
Impaired communication	15.0 (5.0)	14.0 (5.1)
Stereotyped behaviours	5.2 (3.3)	5.9 (2.8)

[1]AD: Autistic Disorder; [2]PDDNOS/AS: PDD-NOS or Asperger Syndrome.

reliably than past ratings, especially since many participants were aged considerably above 4–5 years.

Model estimations and model comparisons

The default maximum likelihood (ML) estimation procedure, as implemented in M*plus* (Muthén & Muthén, 1998–2007, Version 4.2), could have been applied to estimate the models and to examine their goodness of fit. Although the kurtosis and skewness values of the 12 ADI-R indicators were not extreme for ML estimation (Boomsma & Hoogland, 2001), a robust ML estimation procedure was preferred to analyse sample covariance matrices **S**. In M*plus*, this procedure is labelled as MLM: maximum likelihood parameter estimation with standard errors and an adjusted model fit statistics that are robust against non-normality. The global fit of the models was assessed using the scaled or mean-adjusted chi-square statistics, χ^2_{SB} of Satorra and Bentler (1994), along with other fit indices, the values of which were roughly evaluated using cut-off criteria formulated by Hu and Bentler (1999); cf. Marsh, Hau, and Wen (2001).

To determine whether the three dichotomised covariates had substantive effects on the model structure and estimates, the model fit and parameter estimates of the selected groups were compared in a stepwise procedure (Brown, 2006). The primary objective was to investigate whether main features of the model are invariant across the two groups selected for each of the three covariates. To that purpose, it was investigated whether the model structure has the same form in each of the groups. If that was the case, the next step was to check whether the factor loadings are the same across the groups.

More specifically, for each of the three covariates, two multiple-group analyses were applied to examine the difference in chi-square fit between Model B2 with invariance restrictions on the factor loadings (i.e., factor loadings were assumed to be equal in the two selected groups) and Model B2 with no parameter restrictions but an equal structural form (i.e., the unrestricted Model B2). When the value of the estimated chi-square difference test, denoted as $\chi^2(\Delta)$, is not significant, this indicates that the hypothesis of invariant factor loadings should not be rejected. The estimated chi-square difference test was calculated as follows: $\chi^2(\Delta) = (\chi^2_{ML\text{-}0} - \chi^2_{ML\text{-}1})/c(\Delta)$, where $\chi^2_{ML\text{-}0}$ and $\chi^2_{ML\text{-}1}$ are the regular ML chi-square fit estimates of the restricted Model B2 with equal factor loadings and the unrestricted Model B2, respectively, and $c(\Delta)$ is the scaling correction factor for the difference test, i.e., a function of degrees of freedom and scaling correction factor $c = \chi^2_{ML}/\chi^2_{SB}$ (see Muthén & Muthén, 2007; Brown, 2006).

Results

Cross-validation

DSM model. The two separate DSM models (with ratings related to *current* or *past status*) could not be estimated, because both had an empirical identification problem: an estimated correlation larger than one between the latent factors *S* (impaired socialisation) and *C* (impaired communication). The same problem in estimating the DSM model occurred in our previous study, where it was discussed in detail.

Model B2. The goodness-of-fit estimates for Model B2, and for the one- and two-factor model, with ADI-R ratings related to both the *current* and *past status* are presented in Table 2. The results indicate that, compared to the other two models, Model B2 has the best fit regarding all discrepancy measures. Although the value of χ^2_{SB} was too high, the RMSEA, SRMR, TLI and CFI estimates of Model B2 with *current* and *past ratings* were quite good.

The completely standardised loadings of Model B2 with ADI-R *current ratings* are: *F1* (.72, .66, .74, .70, .64), *F2* (.32, .47, .43, .59, .67), and *F3* (.96, .86). The estimated correlations between the three factors are moderate: $r(F1, F2) = .28$, $r(F1, F3) = .32$, and $r(F2, F3) = .45$. The factor determinacy scores of the three factors *F1*, *F2*, and *F3*, which provide an estimate of the squared multiple correlation between each factor and the indicators (see Grice, 2001), are .91, .83 and .97, respectively. This is quite satisfying, because wildly different rankings on factor scores become unlikely.

Table 2 Goodness-of-fit values for the cross-validation of Model B2, compared to those of a one- and a two-factor model with data from Utrecht (*N = 263*; ADI-R *current* and *past status*)

	Global fit measures					MLM-based statistics					
	df	χ^2_{ML}	*p*	χ^2_{SB}	*p*	RMSEA	SRMR	TLI	CFI	AIC	ECVI[1]
ADI-R current status											
Model B2	51	83.5	.00	83.4	.00	.05	.05	.96	.97	10924	.52
One-factor model	54	486.4	.00	483.0	.00	.17	.12	.44	.54	11321	1.95
Two-factor model	53	444.3	.00	445.2	.00	.17	.11	.47	.58	11281	1.67
ADI-R past status											
Model B2	51	116.4	.00	118.2	.00	.07	.06	.91	.93	11244	.67
One-factor model	54	270.3	.00	272.4	.00	.12	.09	.74	.78	11392	1.44
Two-factor model	53	209.4	.00	210.7	.00	.11	.08	.81	.84	11333	1.01

Note: RMSEA = root mean square error of approximation; SRMR = standardised root mean square residual; TLI = Tucker–Lewis Index; CFI = Comparative Fit Index; AIC = Akaike's Information Criterion; [1]ECVI = Expected Cross Validation Index and these values were computed using version 8.8 of the LISREL program (Jöreskog & Sörbom, 1996).

Table 3 Goodness-of-fit values for Model B2 for groups of participants: a) AD versus PDD-NOS/AS, b) low versus high FIQ, and c) children versus adolescents; combined data from Utrecht and Groningen (*N = 356*; ADI-R *current status*)

	Global fit measures					MLM-based statistics					
Model B2	*df*	χ^2_{ML}	*p*	χ^2_{SB}	*p*	RMSEA	SRMR	TLI	CFI	AIC	ECVI
ASD diagnosis											
Autistic Disorder (*N = 143*)	51	69.0	.05	70.3	.04	.05	.06	.96	.96	6151	.88
PDD-NOS/AS (*N = 213*)	51	70.3	.04	69.9	.04	.04	.05	.96	.97	8760	.58
IQ level											
Low (FIQ < 85; *N = 136*)	51	69.7	.04	67.8	.06	.05	.06	.95	.96	5795	.88
High (FIQ ≥ 85; *N = 220*)	51	71.9	.03	72.3	.03	.04	.05	.97	.97	9151	.56
Age											
Child (4–12 yrs; *N = 247*)	51	88.9	.00	88.7	.00	.05	.05	.94	.96	10501	.57
Adolescent (13–24 yrs; *N = 108*)	51	65.1	.09	63.9	.11	.05	.07	.91	.93	4059	1.06

Table 4 Goodness-of-fit values under model invariance testing for Model B2 for groups of participants: a) AD versus PDD-NOS/AS, b) low versus high FIQ, and c) children versus adolescents; combined data from Utrecht and Groningen (*N = 356*; ADI-R *current status*)

	Global fit measures										MLM-based statistics			
Model B2	*df*	χ^2_{ML}	*p*	χ^2_{SB}	*p*	*c*	*c*(Δ)	χ^2(Δ)	*df*(Δ)	*p*	RMSEA	SRMR	TLI	CFI
ASD diagnosis (*N = 143, N = 213*)							.97	14.6	8	.07				
Equal form	103	139.3	.01	140.2	.00	.99					.05	.05	.96	.97
Equal loadings	111	153.4	.00	154.7	.00	.99					.05	.06	.96	.96
IQ level (*N = 136, N = 220*)							.90	9.5	9	.39				
Equal form	102	141.6	.01	140.1	.01	1.01					.05	.05	.96	.97
Equal loadings	111	150.2	.01	149.9	.01	1.02					.04	.06	.96	.97
Age (*N = 247, N = 108*)							1.01	7.0	9	.63				
Equal form	102	154.0	.00	152.4	.00	1.01					.05	.04	.93	.95
Equal loadings	111	161.1	.00	159.5	.00	1.01					.05	.06	.94	.95

Covariate group comparisons In Tables 3 and 4 the results of the multiple-group analyses are shown. It should be noted that an empirical identification problem occurred while analysing the equal form model given the sample covariance matrix of the PDD-NOS/AS group: the estimated residual or error variance of *S2* was –.003, a marginal Heywood case. An admissible solution was obtained by fixing that variance to the value of +.001: an artificial 'removal' of the problem. Hence, results for the ASD comparisons need to be evaluated with caution. Further reflections on this issue follow in the discussion.

As shown in Table 3, there are no large differences in model fit indices between the pairwise selected groups. The findings of the multiple-group analyses shown in Table 4 demonstrate that the chi-square difference test result was not significant regarding the group comparisons for all three covariates: symptom severity [$\chi^2(\Delta) = 14.6$; $df(\Delta) = 8$; $p = .07$], IQ level [$\chi^2(\Delta) = 9.5$; $df(\Delta) = 9$; $p = .39$], and age [$\chi^2(\Delta) = 7.0$; $df(\Delta) = 9$; $p = .63$]. The equality of factor loadings is most clearly demonstrated between the age groups and less between the two ASD diagnosis groups. The combined results of model fit and parameter estimates show that Model B2 is fairly invariant, certainly with regard to the two intelligence and age levels.

Discussion

In this study it was examined a) whether a new symptom model for autism could be cross-validated in an independent sample of ASD children and adolescents, and b) to what extent the symptom model was invariant under dichotomised levels of symptom severity, intelligence, and age.

The new symptom model was previously tested in a sample of 255 verbal and nonverbal children and adolescents, and only half of these participants had a disorder on the autism spectrum (Van Lang et al., 2006). It was therefore interesting to investigate whether this symptom model could be cross-validated in an independent sample of 263 verbal ASD children and adolescents. The goodness-of-fit values of the model were satisfactory and proved again to be much better than those of the one- and two-factor model. In addition, it turned out that the DSM model could not be properly estimated, owing to a similar empirical identification problem found in our previous study: an improper estimated correlation among the symptom domains impaired socialisation and impaired communication. Apparently, a symptom factor that consists of symptoms about impairments in socialisation and communication is necessary. This issue has been raised by many authors (e.g., Robertson et al., 1999; Tadevosyan-Leyfer et al.,

2003), and is described as *impaired social communication* in the present model.

Besides *impaired social communication*, Model B2 consists of the symptom factors *stereotyped features in speech and behaviour* and *impaired play skills*. The latter symptom factor resembles the theory of mind factor described by Robertson et al. (1999) and Tanguay et al. (1998), and its emergence might explain why Kolevzon et al. (2004) found a different familiality of social play, compared to the familiality of the DSM domain impaired communication in which social play is incorporated. However, the symptom factors of our model are not consistent with why Kolevzon and his colleagues found a different familiality of circumscribed interests and preoccupation with parts of objects, compared to the familiality of the DSM domain about stereotyped behaviours, because the symptom factor *stereotyped features in speech and behaviour* consists of all symptoms of this DSM domain, along with stereotypes in language skills.

The new symptom model was shown to be relatively invariant under dichotomised differences in symptom severity, intelligence, and age. The values of the model's goodness-of-fit statistics and indices were quite similar between the dichotomised covariate groups and comparable to those found in our previous study (Van Lang et al., 2006). In addition, the non-significant differences in chi-square values between the comparison groups indicated that the factor loadings were about the same. There was, however, an empirical identification problem in the test of model-form equality in the group of participants with a PDD-NOS/AS diagnosis. This may be due to the fact that the factor *impaired play skills* has two sumscore indicators only, which put restrictions on finding a fully admissible solution for this diagnosis group. The estimated error variance of *S2* was relatively low in all dichotomised covariate groups, and strongest in the PDD-NOS/AS group where it resulted in a Heywood case.

In a study under progress, it is further examined why these estimation problems occurred by looking at associations between ADI-R subdomain scores, or ADI-R item scores, and the three covariates as continuous variables. For instance, the results of a MIMIC (multiple indicators, multiple causes) type of modelling might reveal how the eight ADI-R items incorporated in the two indicators of the symptom factor *impaired play skills* are related to younger or older children, or adolescents. These analyses might thus provide a more detailed description of how separate symptoms, symptom factors, and covariates are associated using cross-sectional data.

Longitudinal data are necessary, however, to examine how symptoms within the symptom factors might change over time. As shown by Charman et al. (2005) and Sigman and McGovern (2005), it may be that within the symptom factor *impaired social communication*, impaired use of nonverbal behaviours (*S1*) will improve over time, whereas impaired initiation and maintenance of conversational interchange (*C3*) remain fairly stable. Changes in the correlations among the symptom factors may also be examined to see whether the symptom factors are relatively separate constructs. Overall, the estimated correlations among the factors were moderate, and within the dichotomised covariate groups, they ranged between .14 and .62, with a median of .34. These correlations are higher than those reported by Ronald et al. (2006a, 2006b) for the DSM triad domains.

In conclusion, our new symptom model, Model B2, may prove to be potentially valid and promising for further development and understanding of an empirically validated model for autism. Recently, Frazier, Youngstrom, Kubu, Sinclair, and Rezai (in press) have tested several new symptom models for autism in a large group of *N = 1,170* verbal individuals, and two models appeared to have the best fit and stability across two age groups: a model similar to Model B2, and a two-factor model that was a combination of Model B2 and the two-factor model examined in this study. In addition, Georgiades et al. (2007) found support for a three-factor symptom model that has similarities with Model B2 (e.g., *impaired social communication*), but differs with respect to *stereotyped features in speech and behaviour* (they found two separate factors for rigidity and repetitiveness in these features) and *impaired play skills* (which they did not find). They were able to estimate the DSM model (*N = 209*), but the DSM model failed to meet acceptable model criteria. These findings reaffirm that the DSM model is empirically not a firm model, and that other new symptom models, like Model B2, may be promising alternatives.

There are two limitations that need further attention. First, it is necessary to obtain more replications of the new symptom model with instruments other than the ADI-R, since the model is now based on ADI-R ratings exclusively. Although the ADI-R is a frequently used semi-structured interview that yields a comprehensive description of autistic behaviours during the age of 4–5 and current age, it would be interesting to investigate whether a comparable symptom model would emerge if, for instance, the Social Communication Questionnaire (Rutter, Bailey, & Pickles, 2003) were used. Second, the symptom model would gain in clinical weight if its symptom structure sustains over time within the same individuals. At present, the symptom model is studied using cross-sectional data only. The clinical validity of the symptom model would improve if different developmental trajectories within the symptom factors could be demonstrated using longitudinal data.

Acknowledgements

Data were obtained from studies that were supported by grants from the Netherlands Organisation for Health Research and Development (ZON/MW) and from the Korczak Foundation for autism and related disorders.

Correspondence to

Natasja D.J. van Lang, Leiden University Medical Center, Curium, Department of Child and Adolescents Psychiatry, PO Box 15, 2300 AA Leiden, The Netherlands; Tel: +31 71 5159633; Email: N.D.J.van_Lang@Curium.nl

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (edn 4 text-revision (DSM-IV-TR)). Washington, DC: American Psychiatric Association.

Bleichrodt, N., Resing, W.C.M., Drenth, P.J.D., & Zaal, J.N. (1987). *Intelligentie-meting bij kinderen. Empirische en methodologische verantwoording van de Gereviseerde Kinder Intelligentie Test* [Intelligence measurement in children. Empirical and methodological justification of the Revised Child Intelligence Test]. Lisse, The Netherlands: Swets & Zeitlinger.

Boomsma, A., & Hoogland, J.J. (2001). The robustness of LISREL modeling revisited. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural equation modeling: Present and future. A Festschrift in honor of Karl Jöreskog* (pp. 139–168). Lincolnwood, IL: Scientific Software International.

Brown, T.A. (2006). *Confirmatory factor analysis for applied research*. New York: The Guilford Press.

Charman, T., Taylor, E., Drew, A., Cockerill, H., Brown, J.A., & Baird, G. (2005). Outcome at 7 years of children diagnosed with autism at age 2: Predictive validity of assessments conducted at 2 and 3 years of age and pattern of symptom change over time. *Journal of Child Psychology and Psychiatry*, *46*, 500–513.

Coplan, J., & Jawad, A.F. (2005). Modeling clinical outcome of children with autistic spectrum disorders. *Pediatrics*, *116*, 117–122.

De Bildt, A., Sytema, S., Kraijer, D., & Minderaa, R. (2005). Prevalence of pervasive developmental disorders in children and adolescents with mental retardation. *Journal of Child Psychology and Psychiatry*, *46*, 275–286.

Frazier, T.W., Youngstrom, E.A., Kubu, C.S., Sinclair, L., & Rezai, A. (in press). Exploratory and confirmatory factor analysis of the Autism Diagnostic Interview-Revised. *Journal of Autism and Developmental Disorders*.

Georgiades, S., Szatmari, P., Zwaigenbaum, L., Duku, E., Bryson, S., Roberts, W., Goldberg, J., & Mahoney, W. (2007). Structure of the autism symptom phenotype: A proposed multidimensional model. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*, 188–196.

Grice, J.W. (2001). Computing and evaluating factor scores. *Psychological Methods*, *6*, 430–450.

Hu, L., & Bentler, P.M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Convential criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.

International Molecular Genetic Study of Autism Consortium (IMGSAC). (2001). Further characterization of the autism susceptibility locus AUTS1 on chromosome 7q. *Human Molecular Genetics*, *10*, 973–982.

Jöreskog, K.G., & Sörbom, D. (1996). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Chicago, IL: Scientific Software International.

Kolevzon, A., Smith, C.J., Schmeidler, J., Buxbaum, J.D., & Silverman, J.M. (2004). Familial symptom domains in monozygotic siblings with autism. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, *129B*, 76–81.

Lord, C., Risi, S., Lambrecht, L., Cook, E.H., Jr, Leventhal, B.L., DiLavre, P.C., Pickles, A., & Rutter, M. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*, 205–223.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*, 659–685.

Marsh, H.W., Hau, K.T., & Wen, Z. (2001). In search of golden rules: Comment on hypothesis testing approaches to setting cut-off values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*, 320–341.

McGovern, C.W., & Sigman, M. (2005). Continuity and change from early childhood to adolescence in autism. *Journal of Child Psychology and Psychiatry*, *46*, 401–408.

Muthén, L., & Muthén, B. (1998–2007). *Mplus user's guide*. Los Angeles, CA: Muthén and Muthén.

Muthén, L., & Muthén, B. (2007). *Chi-square difference testing using the Satorra–Bentler scaled chi-square*. Available at http://www.statmodel.com/chidiff.html.

Robertson, J.M., Tanguay, P.E., L'Ecuyer, S., Sims, A., & Waltrip, C. (1999). Domains of social communication handicap in autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *38*, 738–745.

Ronald, A., Happé, F., Bolton, P., Butcher, L.M., Price, T.S., Wheelwright, S., Baron-Cohen, S., & Plomin, R. (2006a). Genetic heterogeneity between the three components of the autism spectrum: A twin study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*, 691–699.

Ronald, A., Happé, F., Price, T.S., Baron-Cohen, S., & Plomin, R. (2006b). Phenotypic and genetic overlap between autistic traits at the extremes of the general population. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*, 1206–1214.

Rutter, M., Bailey, A., & Pickles, A. (2003). *Social Communication Questionnaire (SCQ)*. Los Angeles: Western Psychological Services.

Satorra, A., & Bentler, P.M. (1994). Correlations to test statistics and standard errors in covariance structure analysis. In A. von Eye, & C.C. Clogg (Eds.), *Latent variable analysis: Applications for developmental research* (*pp*. 300–419). Thousand Oaks, CA: Sage.

Sigman, M., & McGovern, C.W. (2005). Improvement in cognitive and language skills from preschool to adolescence In autism. *Journal of Autism and Developmental Disorders*, *35*, 15–23.

Tadevosyan-Leyfer, O., Dowd, M., Mankoski, R., Winklosky, B., Putnam, S., McGrath, L., Tager-Flusberg, H., & Folstein, S. (2003). A principal components analysis of the Autism Diagnostic Interview-Revised. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*, 864–872.

Tanguay, P.E., Robertson, J., & Derrick, A. (1998). A dimensional classification of autism spectrum disorder by social communication domains. *Journal of the American Academy of Child and Adolescent Psychiatry*, *37*, 271–277.

Van Bon, W.H.J. (1984). *Raven's coloured progressive matrices: Nederlandse normen en enige andere uitkomsten van onderzoek* [Dutch norms and other study results]. Lisse, The Netherlands: Swets & Zeitlinger.

Van der Steene, G., Haasen, P.P., De Bruyn, E.E.J., Coetsier, P., Pijl, Y.J., Poortinga, Y.H., Lutje Spelberg, H.C., Spoelders-Claes, R., & Stinissen, J. (1986). *Wechsler Intelligence Scale for Children-Revised* [*Dutch version*] (*WISC-RN*). Lisse, The Netherlands: Swets & Zeitlinger.

Van Lang, N.D.J., Boomsma, A., Sytema, S., de Bildt, A.A., Kraijer, D.W., Ketelaars, C., & Minderaa, R. (2006). Structural equation analysis of a hypothesised symptom model in the autism spectrum. *Journal of Child Psychology and Psychiatry*, *47*, 37–44.

Manuscript accepted 8 January 2008